A Hybrid Model for Secure Data Transfer in Audio Signals using HCNN and DD DWT

B. Geetha vani [1]
Research scholar, Dept. of CSE,
JNTU Kakinada. AP, India

Prof. E. V. Prasad [2]
Professor, Dept. of CSE & Rector,
JNTU Kakinada, AP, India

***Abstract*—In today's world, there are a number of cryptographic and steganography techniques used in order to have secured data transfer between a sender and a receiver. In this paper a new hybrid approach that integrates the merits of cryptography and audio steganography is presented. First, the original message is encrypted using chaotic neural network and the resultant cipher text is embedded into a cover audio using Double Density Discrete Wavelet Transform (DD DWT). The resultant stego audio is transmitted to the receiver and the reverse process is done in order to get back the original plain text. The proposed method presents a Steganography scheme along with the cryptography scheme which enhances the security of the algorithm.**

Keywords—Cryptography; Hopfield Chaotic Neural Network; Audio Steganography; Double Density Discrete Wavelet Transform.

I. INTRODUCTION

The increasing internet usage stems from the growing availability of global communication technology that has led to electronically included information gathering and distribution. However, the challenge it presents in terms of information security is enormous. The need to secure information within the global network is of paramount importance so that the user information is preserved until it reaches its destination undisclosed. Providing a secure framework that conceals information content and sender/receiver identity is a matter of prime interest.

The two popular approaches to information confidentiality are Cryptography and Steganography [1,2,3]. Cryptography is the study and practice of protecting information by data encoding and transformation techniques. Steganography, a concealed writing, is the art and science of hiding the fact that communication is taking place.

Steganography techniques, based on the cover file, can be categorized as Image steganography, Text Steganography, Audio steganography and Video steganography. In digital audio steganography system, secret message is embedded in audio file. The binary sequence of audio file is slightly changed by adding secret message to it. This modification

Should not be made available to the human ear. Embedding secret messages in audio file is more difficult than embedding information in digital image. In order to hide secret messages, various methods for embedding information in digital audio have been introduced. These methods range from simple techniques that insert information in the form of noise in audio signal to more powerful methodologies using signal processing techniques. Many audio steganography methods use Least Significant Bit (LSB) insertion technique [4,6] to hide the secret message. But techniques have been developed to detect secret message which is present at LSB position [12]. Hence, an improvement over this is use of robust audio steganography techniques using wavelets [11, 16, 17].

In the proposed system, audio steganography method uses double density discrete wavelet transforms. In order to provide better security chaotic neural network is used for encrypting the secret message. The rest of the paper is organized as follows. In Section 2, Literature review is presented. In Section 3, the proposed system is described. In Section 4, experimental results of proposed approach are shown. Concluding remarks are provided in section 5.

II. LITERATURE REVIEW

A. Audio Steganography Techniques

Audio Steganography techniques can be adopted either in temporal domain or transform domain.

1) Temporal domain Techniques

a) LSB

LSB is one of the earliest, simplest and commonly used methods, for hiding information in audio steganography. In LSB method, as shown in Fig.1, the least significant bits of the cover media/original audio is altered to include the secret message. Even though this is a simple method, an attacker can easily extract the secret message from the stego object.

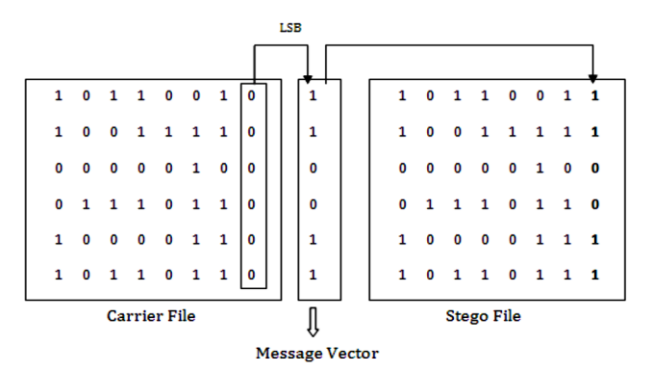

Fig. 1. LSB Insertion

b) Parity coding

Parity coding technique operates on a group of samples instead of individual sample. Here individual samples are grouped and parity of each group is calculated. For inserting message bit one by one, check the parity bit of a group of samples. If the parity bit and message bit matches then no operation need to be performed else change the LSB's of any one of the individual samples in that group to make the parity bit equal to the message bit.

c) Echo hiding

In echo hiding method data is embedded in the echo part of the host audio signal. The echo is a resonance added to the host signal and hence the problem with the additive noise is avoided here. While using echo hiding, to avoid echo audibility, three important parameters to be considered are initial amplitude, offset (delay), and decay rate. The method suffers from lenient detection and low detection ratio. Due to its low embedding rate and low security use of this technique is not interesting among researchers.

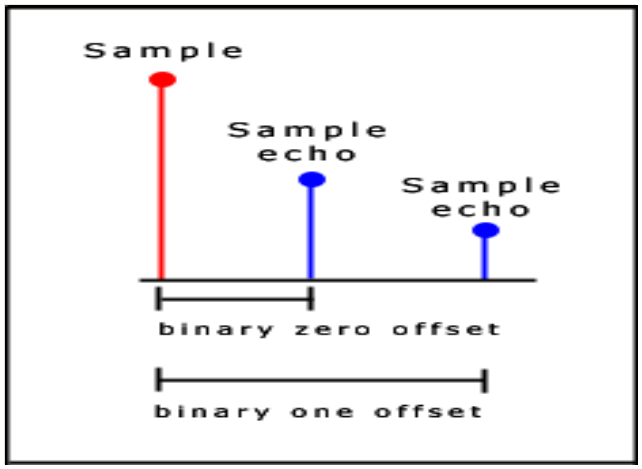

Fig. 2. Echo Hiding

B. Transform domain Techniques

Frequency domain techniques and wavelet domain technique comes under transform domain. The main techniques that can usually adopted in frequency domain include tone insertion, phase coding and spread spectrum technique.

1) Tone insertion

Frequency masking property is exploited in tone insertion method. A weak pure tone is masked in the presence of a stronger tone. This property of inaudibility is used in different ways to embed information.

2) Phase coding

Phase coding method is based on the fact that the phase components are not audible to human as noise components. This method as shown in Fig. 3, embeds the secret message bits as phase shift in the phase spectrum of the original audio signal. The method tolerates better signal distortion, better robustness, but it does not survive low pass filtering.

Here the secret message is inserted only at the phase vector of the first signal segment.

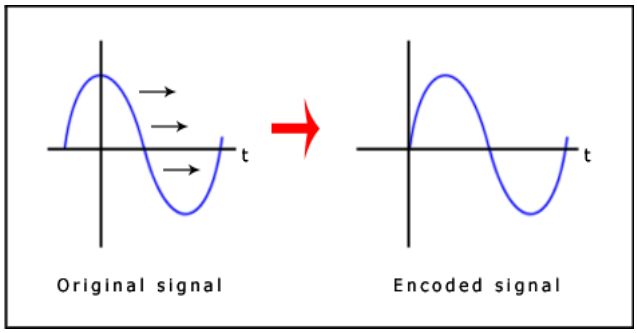

Fig. 3. Phase Coding

3) Spread spectrum technique

This technique, explained in Fig. 4, takes the advantage of masking property of Human Audio System (HAS). A masking threshold is calculated using a psycho-acoustic model. The spread signal now lies below the masking threshold. Apart from phase shifting, here the secret message is distributed along with the host signal. Here the final signal occupies a bandwidth which is more than what is actually required for transmission.

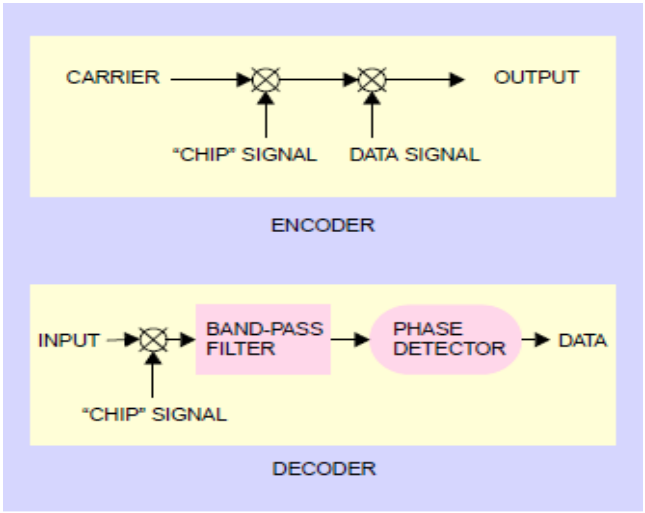

Fig. 4. Spread Spectrum

4) Wavelet domain

Wavelet domain technique, that uses wavelet coefficients, is suitable for frequency analysis because of its multi-resolution property and provides access to both most significant parts and details of spectrum. With the use of inverse transform, the stego signal can be reconstructed.

III. Proposed System

In the Proposed hybrid model, Steganography is combined with Cryptography to transmit message in a highly secured manner and makes the system theoretically and practically unbreakable. For embedding the information, the steps involved are as follows,

1) Get the Plain text which is to be sent to the recipient.

2) Transform the plaintext in to cipher text by applying an encryption process using chaotic neural network.

3) Embed the cipher text inside the cover audio file using double density discrete wavelet transform.

4) The resulted stego audio file is communicated through any communication channel to the receiver.

For extraction of information, the steps involved at the receiver side are as follows,

1) On receiver side Median noise filter is applied to remove noise from the stego audio.

2) By applying inverse DD DWT the embedded message is extracted from the Stego audio file.

3) Since the obtained message is in the scrambled form, decryption is performed.

4) Finally, the receiver is able to read the actual secret message sent at the sender's end.

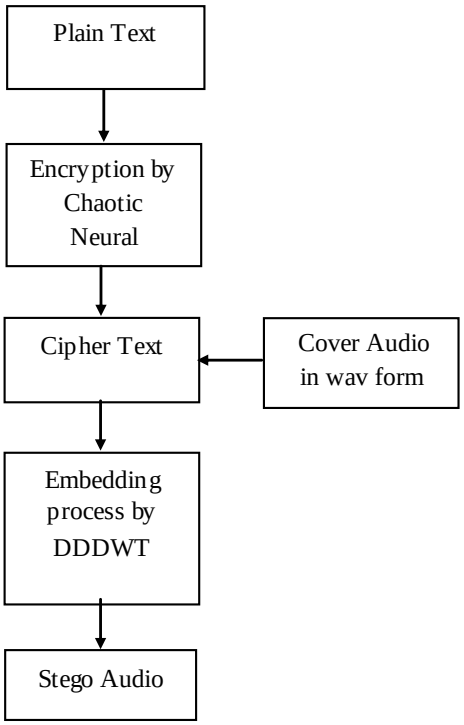

Fig. 5. Embedding process

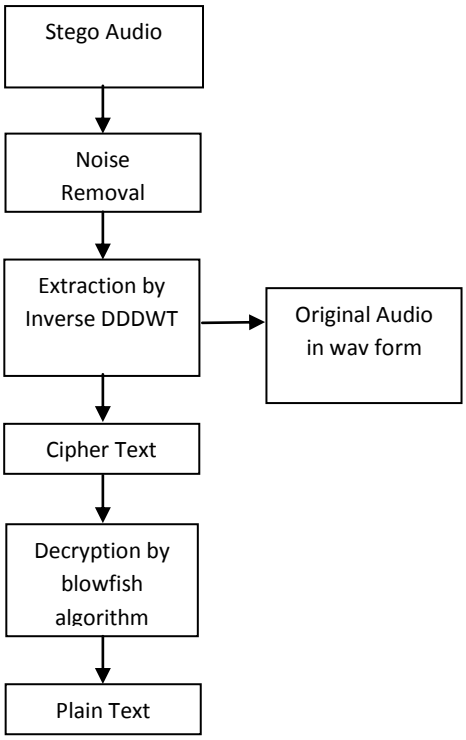

Fig. 6. Extraction process

A. Cryptographic Approach

1) Hopfield Chaotic Neural Network Based Encryption

The encryption methodology adopted for encrypting text characters plays a vital role in deciding the embedding capacity and the level of robustness and security of the entire Steganography system. Hopfield Chaotic Neural network is a suitable environment for cryptography because of some interesting properties like ergodicity, sensitive dependence of initial conditions and control parameters and high speed of information transmission. Yu et al [7] designed a delayed chaotic neural network based cryptosystem, which makes use of the chaotic trajectories of two neurons to generate basic binary sequences for encrypting plaintext. In Chaotic Neural Network, the weights and biases are determined by a chaotic sequence, a binary random deterministic sequence, and are used to mask or to scramble the original information. The encryption algorithm [8] is used for obtaining the cipher text. The Chaotic neural network consumes less computational power and the sequence generated using this is unpredictable leading to highly secured and efficient in terms of power.

2) Double-Density Discrete Wavelet Transform

The double-density DWT is an improved critically sampled DWT with following additional properties:

a) It employs one scaling function and two distinct wavelets, which are designed to be offset from one another by one half.

b) The double-density DWT is over complete by a factor of two and

c) It is nearly shift-invariant.

In two dimensions, this transform outperforms the standard DWT in terms of denoising. However, there is room for improvement because not all of the wavelets are directional. That is, although the double-density DWT utilizes more wavelets, some lack a dominant spatial orientation, which prevents them from being able to isolate those directions.

B. Implementation of DD-DWT

To implement the double-density DWT, it is necessary to first select an appropriate filter bank structure. The filter bank proposed in Fig. 7 illustrates the basic design of the double-density DWT.

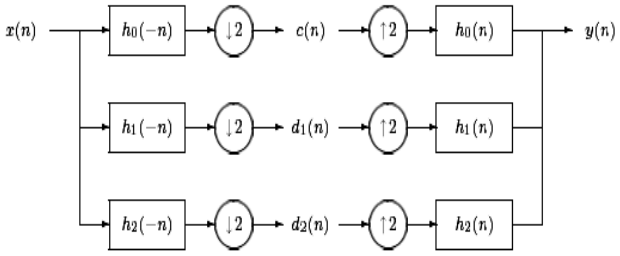

Fig. 7. A 3-Channel Perfect Reconstruction Filter Bank.

The analysis filter bank consists of three analysis filters - one low - pass filter denoted by $h_0(-n)$ and two distinct high-pass filters denoted by $h_1(-n)$ and $h_2(-n)$. As the input signal x(n) travels through the system, the analysis filter bank decomposes it into three sub-bands, each of which is then down-sampled by 2. From this process the signals c (n), $d_1(n)$, and $d_2(n)$, which represent the low frequency (or coarse) sub-band, and the two high frequency (or detail) sub-bands are obtained.

The synthesis filter bank consists of three synthesis filters - one low-pass filter denoted by $h_0(n)$ and two distinct high-pass filters denoted by $h_1(n)$ and $h_2(n)$ - which are essentially the inverse of the analysis filters. As the three sub-band signals travel through the system, they are up-sampled by two, filtered, and then combined to form the output signal y (n).

C. Significance of the Hybrid Model

The proposed method integrates two different techniques for the secured data transmission. They are

- Enciphering & Deciphering phase with the Cryptography
- Embedding & Extraction of data with the Steganography

1) It's quite hard for the eavesdroppers to realize the chaotic neural network encryption hence probability of attack is less when compared with the normal encryption algorithms.

2) Embedding process is done using DD DWT, such that the resulting stego audio is similar to original audio and provides robustness.

3) In this model on the receiver side median filter is used to remove noises in stego audio.

IV. Results And Analysis

Experiments have been conducted to prove the efficiency of the proposed algorithm. The Quantitative performance of the proposed algorithm is evaluated based on Peak signal to noise ratio (PSNR) and Mean Square Error (MSE)which are given in equations 1 & 2 respectively.

The peak signal to noise ratio (PSNR) was utilized to evaluate the stego audio quality. PSNR [15] is often expressed on a logarithmic scale in decibels (dB), it is defined as:

$$PSNR = 10 * \log_{10}(255^2/MSE)\ (dB) \qquad (1)$$

Where MSE (15) is the mean square error between the cover and stego audio. For a cover audio whose size is defined in terms of W and H, MSE is defined as:

$$MSE = \frac{1}{W*H}\left(\sum_{i=1}^{W}\sum_{j=1}^{H}(A_{ij}\text{-}A^{1}_{ij})^2\right) \qquad (2)$$

Where W is the amplitude of the signal, H is the frequency of the audio signal and A_{ij} and A^{1}_{ij} are pitch values of cover and Stego audio.

An audio file with ".wav" extension has been selected as cover file. Modification of bits should not degrade sound quality. Figure 8 shows graph of original audio which is used as host file. Figure 9 shows graph of audio after embedding and figure 10 shows graph of recovered audio after extraction. Graph of original audio, embedded audio and recovered audio is nearly same. These graphs are plotted Sample Numbers versus amplitude. The simulation was carried out in MATLAB R2010b software.

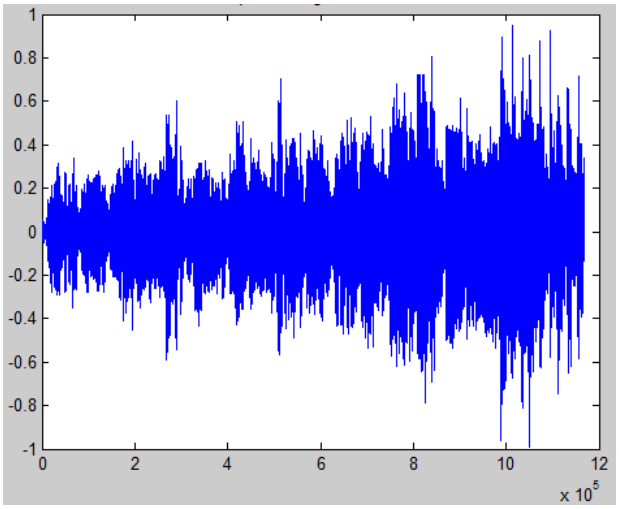

Fig. 8. Original Audio.

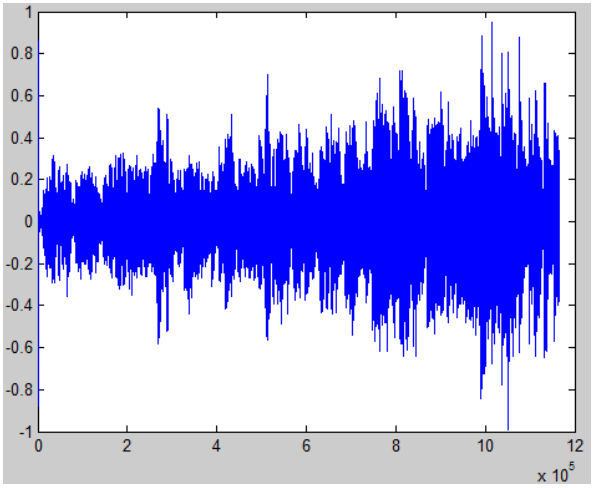

Fig. 9. Stego Audio

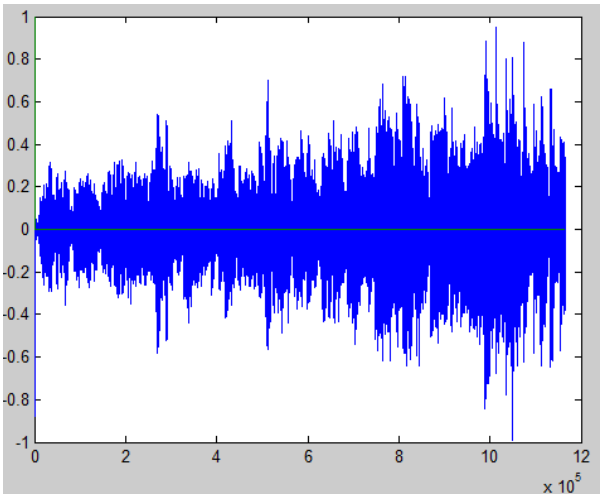

Fig. 10. Recovered Audio

Fig. 11 is Original Message and Fig. 12 is Recovered Message. These two messages are 100% similar.

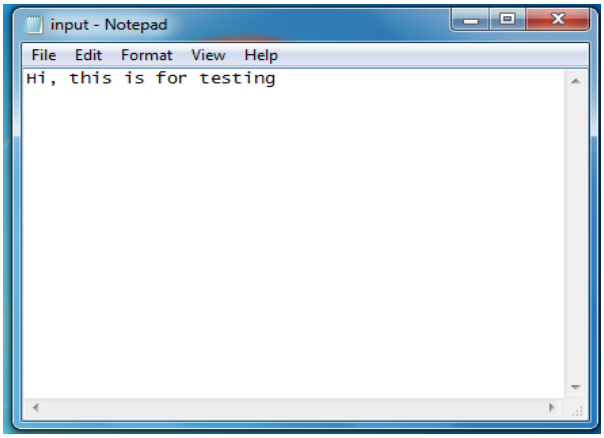

Fig. 11. Original text message

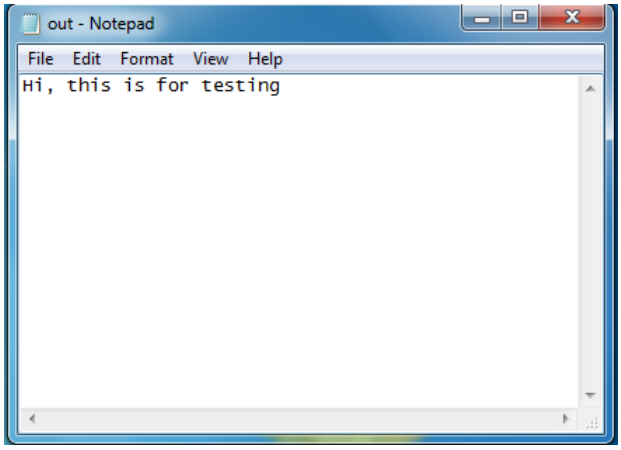

Fig. 12. Recovered message

TABLE I. MSE values of different audio files for different message sizes

Message size in bytes	Male	Female	Male song	Female song
10	$2.24*10^{-5}$	$1.75*10^{-5}$	$5.01*10^{-6}$	$6.8*10^{-6}$
20	$5.05*10^{-5}$	$3.94*10^{-5}$	$1.12*10^{-5}$	$1.54*10^{-5}$
33	$8.97*10^{-5}$	$7.00*10^{-5}$	$2.00*10^{-5}$	$2.74*10^{-5}$
57	$1.40*10^{-4}$	$1.09*10^{-4}$	$3.13*10^{-5}$	$4.28*10^{-5}$

TABLE II. PSNR values of different audio files for different message sizes

Message size in bytes	Male	Female	Male song	Female song
10	70.55	71.63	77.05	75.70
20	67.03	68.10	73.53	72.18
33	64.53	65.61	71.03	69.68
57	62.59	63.67	69.10	67.74

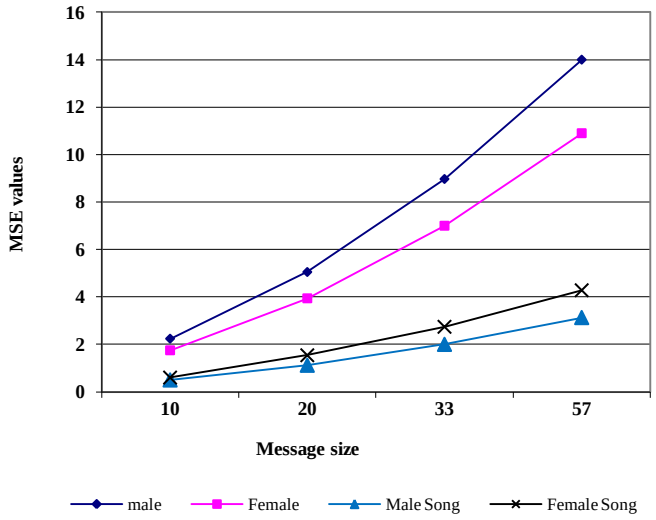

Fig. 13. Graphical representation of MSE values for Different Songs

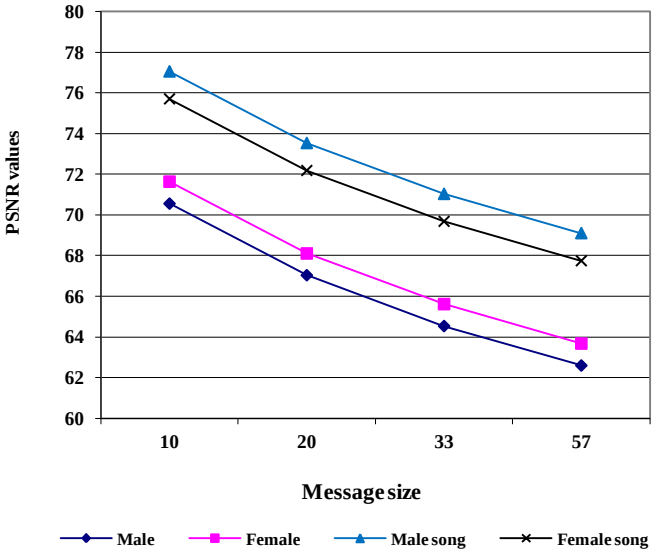

Fig. 14. Graphical representation of PSNR values for different songs

TABLE III. MSE values of different categories for Audio file with same text content

Audio file	MSE
Hip-hop	$4.7*10^{-6}$
Jazz	$4.4*10^{-6}$
Pop	$4.6*10^{-6}$
Rock	$4.5*10^{-6}$

TABLE IV. PSNR values for different categories of audio file with same text content

Audio file	PSNR
Hip-hop	77.31
Jazz	77.56
Pop	77.36
Rock	77.51

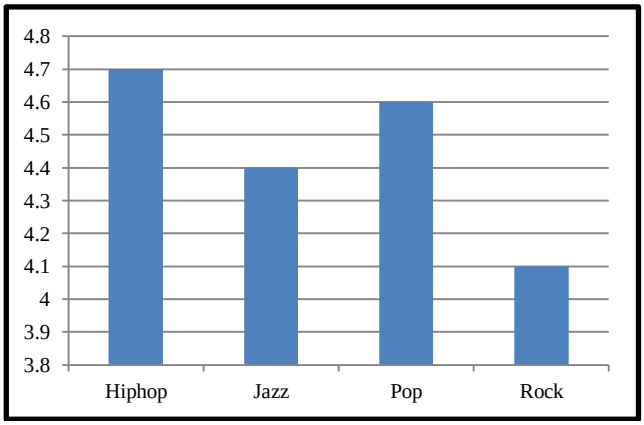

Fig. 15. Graphical representation of MSE values for Different music files

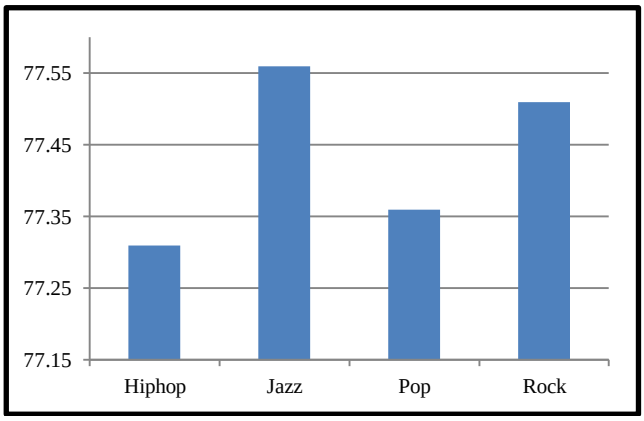

Fig. 16. Graphical representation of PSNR values for Different Music files

TABLE V. CPU time taken for encryption by CNN and with Blowfish algorithm for text files of different sizes

Text Size (in bytes)	CNN algorithm (in seconds)	Blowfish algorithm (in seconds)
50	0.0156	2.0156
100	0.0313	2.0756
150	0.0469	2.1044
200	0.0625	2.1406

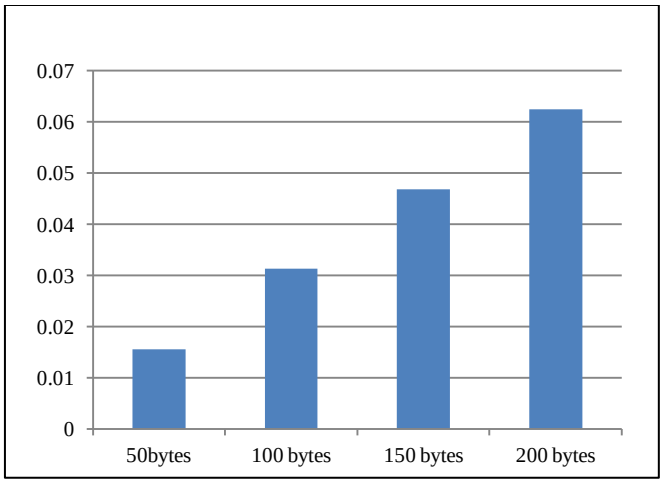

Fig. 17. Graphical representation of CPU time with CNN encryption algorithm

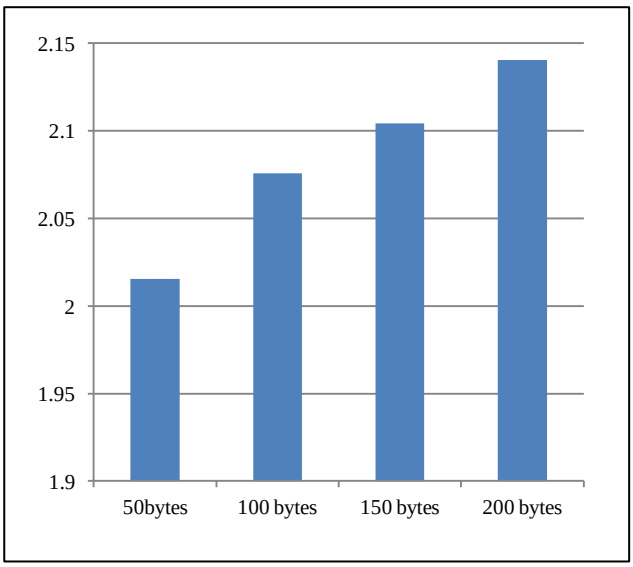

Fig. 18. Graphical representation of CPU time with Blowfish encryption algorithm

V. CONCLUSION

In this Paper, a novel method of Audio Steganography algorithm that uses Double Density Discrete Wavelet Transforms is presented. For providing better security, chaotic neural network encryption scheme is included. The qualitative performance of the proposed system is analyzed. Various sample audio files and music files are considered in .wav format and the MSE and PSNR values obtained after embedding the data have been recorded. The PSNR, MSE varies depending on the amount of data embedded in the audio file and the size of the audio file and better PSNR and low MSE values are obtained with the proposed algorithm. Also it can be observed that CNN encryption algorithm takes less time for the encryption process and is secure than the Blowfish algorithm. The Proposed system shows better performance in terms of both capacity and security. In future, this work can be extended to video data.

REFERENCES

[1] Petitcolas F.A., Anderson R.J., Kuhn M.G, "Information Hiding – A Survey" IEEE. Special Issued on Protection of Multimedia content, 1062-1078 July, 1999.

[2] Katzenbeisser.S, Peticoats.F. "Information hiding Techniques for Steganography and Digital watermarking", Artech House Inc. 2000.

[3] L.Driskell, "Wavelet-based Steganography," Crypto - logia, vol. 28, no. 2, pp.157-174, 2004.

[4] Tian.H, "A Covert Communication Model Based on Least Significant Bits Steganography in voice over IP", Proceedings of 9th International Conference for young computer scientists, IEEE computer Society, pp. 647-652, 2008.

[5] Deng.K, Tian.Y, Yu.X, Yang.Y "A unified block and stream cipher based file encryption" Journal of Global Research in Computer Science, vol.2, No.7, pp.53-57, 2011.

[6] M.Wakiyma, Y.Hidaka, K.Nozaki "A Steganography by low bit coding method with wave files", Sixth International Conference on Intelligent Infomation Hiding and multimedia signal processing, pp.530-533, Oct 2010.

[7] Yu W, Cao J. "Cryptography based on delayed neural networks". PhysLetter A ; 356:333–8. 2006

[8] B.Geethavani, E.V.Prasad. "High Secure Image Steganography Based on Chaotic Neural Network". IJCSNS, Volume13, No.3, pp1-6, March 2013

[9] A.Delforouzi , M.Pooyan, "Adaptive digital audio Steganography based on integer wavelet Transform" Circuits, Systems & Signal Processing, Vol.27, pp.247-259, Mar 2008.

[10] Atoum.M.S, Rababah.O.A , Al.Athili, New technique for hiding data in audio file". Journal of Computer Science, pp.173-177, Apr 2011.

[11] Jisna Antony, Sobin c. Sherly A. P "Audio Steganography in Wavelet Domain – A Survey" International Journal of Computer Applications, Vol.52, No.13, Aug. 2012.

[12] Andrew.D.Ker, "Steganalysis of Embedding in two least significant bits" IEEE Transactions on Information forensics and Security, Vol 2, No.1, Mar 2007.

[13] Mohammad Saleem, Mamoum Suleman, Subariah Ibrahim,"A Steganography method based on hidng secret data in MPEG Audio layer III", International Journal of Computer Science and Network Security, Vol.2, No.5, pp. 184-188, May 2011.

[14] Siwar Rekik, Driss Guerchi, Sid-Ahmed Selouani , Habib Hamam "Speech steganography using wavelet and Fourier Transforms" Rekik et al. EURASIP Journal on Audio, Speech, and Music Processing 2012.

[15] Marghny H Mohamed, Naziha M AL-Aidroos, Mohamed A Bamatraf "A combined image Steganography technique based on edge concept and dynamic LSB", International journal of Engineering Research and Technology,Vol.1 No.8,Oct 2012.

[16] Waffaa S. Ahmed , Loay E. George" Audio Hiding Using Wavelet Transform with Amplitude Modulation" Journal of Al-NahrainUniversity,Vol.16,pp.183-188 March 2013.

[17] Mansour Sheikhan, Kazem Asadollahi , Reza Shahnazi "Improvement of Embedding Capacity and Quality of DWT-Based Audio Steganography Systems" World Applied Sciences Journal, Vol.13, pp.507-516, 2011.